A St abilized Sparse-Matrix U-D Square-Root Implementation of a Large-State Extended Kalman Filter

Dale Boggs[†], Michael Ghil[††], and Christian Keppenne[†]

[†]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, CA 91109-8001, USA*

[††]*Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California, CA 90024-1565, USA*

1. introduction and motivation

The full nonlinear Kalman filter (KF) sequential algorithm is, in theory, well-suited to the four-dimensional data assimilation problem in large-scale atmospheric and oceanic problems (Ghil *et al.* 1981, Ghil and Malanotte-Rizzoli 199 1). Soon after Kalman's (1 960) seminal paper, estimation practitioner's and numerical analysts became aware of numerical difficulties inherent in the discrete KF recursion equations. A number of early researchers (Bellantoni and Dodge 1967, Schmidt *et al.* 1968, Leondes 1970) showed that the KF algorithm can be very sensitive to computer round off and that results may cease to be meaningful as time advances-– even in double precision. A frequent symptom is that an error covariance matrix loses its positive definiteness; this is frequently aggravated by true numerical ill-conditioning which cannot be improved by a simple resealing of units. Filter "divergence" (Schlee *et al.* 1967), i.e., numerical results that wildly contradict predicted analytic behavior, encompass at least 3 types of divergence, due to: (1) the effects of computer roundoff, (2) the presence of nonlinearities, and (3) the use of incorrect *a priori* statistics and an erroneous dynamic model.

These computational shortcomings of the original KF algorithm have motivated alternative formulations of the optimal sequential estimator. Still, many recent KF papers seem oblivious to the conventional discrete KF's numerical stability problems, or, at the most, consider these to be a computational nuisance. When implementing the extended KF (EKF) for a nonlinear, primitive-equation model with realistic forcing and topography, this issue cannot be ignored: it is a challenging part of the overall problem. The causes of filter-divergence types (1) and (2) play a prominent role in the design of the Kalman-type estimation algorithms discussed below.

A square-root formulation of the KF has inherently better stability anti numerical accuracy than the conventional Kalman formulation (Jazwinski 1970, Bierman 1977, Kerr 1990). Square-root filters arc algebraically equivalent to the KF, but involve fundamentally different computational methods. A complete suite of square-root type KF recursions was developed primarily for interplanetary spacecraft navigation (Dyer and McReynolds 1969, Thornton and Bierman 1976, Bierman 1977) and yields numerically robust KF implementations. This is, in part, due to the fact that using square-root matrix factors implicitly preserves symmetry and assures nonnegative eigenvalues for the computed covariance; it also reduces the condition number of the matrices that must be inverted.

Square-root algorithms arc, nonetheless, not very widely used, because of the erroneous perception that factorization techniques are too complicated compared to the conventional KF, use too much computer storage, and involve too much computation. This perception is due in large part: (1) to an incomplete understanding of the square-root algorithms that arc heavily dependent on advanced numerical analysis (Golub and Van Loan 1989, Lawson and Hanson 1974), and (2) on the use of inefficient computer implementations.

The optimality of Kalman-type sequential estimators, whether conventional or square-root, comes at the price of an enormous increase in the number of operations for a full EKF implementation. For models with n discrete variables, the cost of advancing the error covariances one time step with the classical Kalman formalism is about n times that of integrating the model itself. Since codes simulating or predicting large-scale flows currently have 105–106 variables, KF implementations have so far been mostly experimental, and in low-resolution models with up to a few thousand variables (Jiang and Ghil 1993, and references therein). To mitigate this computational barrier, a number of banded approximations to the KF (Parrish and Cohn 1985: PC hereafter; Todling and Ghil 1990) have been explored. These approximations to the conventional KF arc based on retaining only those elements of the covariance matrix which differ significantly from zero. Since covariances tend to zero with increasing distance (Balgovind *et al.* 1983), it is feasible to calculate- -and store-- only those diagonals of the forecast-error covariance matrix that contain significant correlations, rather than the entire matrix. The information contained in this greatly reduced set of diagonals is then a good approximation to that contained in the full covariance matrix. These authors also exploited the block-sparseness of the state transition matrix that arises from the finite-ciiffcrcnce scheme; the result was a computationally feasible KF for a linear two-dimensional (2-D) shallow-water system. Unfortunately- within the conventional KF formulation— this banded approximation induces a distinct loss of positive-dcfhitcncss in the propagated covariance matrix after a few assimilation steps, and the method therefore fails.

To achieve a non-diverging EKF that is computationally feasible for large geophysical-flow models and is as close to optimal as possible, wc have extended PC'S banded algorithm, retaining the computation-saving device of diagonal-wise matrix operations on sparse banded matrices, but within a square-root framework. A less restrictive form of the banded approximation was used, to permit the U-D factorization.

2. The U-]) square-root filter

The type of square-mot KF wc implemented, the so-called U-D filter (Bierman 1977), involves a triangular factorization of the covariance matrix $\boldsymbol{P}$ which requires no actual square-root calculations, thus enhancing efficiency. The U-D filter makes use of the matrix decomposition

$$1' = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^T, \tag{1}$$

where $\boldsymbol{D}$ and $\boldsymbol{U}$ arc, respectively, diagonal and unitary upper triangular (n x n) matrices of the form

$$\boldsymbol{D} = diag[d_1, d_2, \cdots, d_n], \qquad \boldsymbol{U} = \begin{bmatrix} 1 & x & x & x & \cdots & x \\ 0 & 1 & x & x & \cdots & x \\ 0 & 0 & 1 & x & \cdots & x \\ 0 & 0 & 0 & 1 & & \\ \vdots & \vdots & \vdots & & \ddots & x \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{2a,b}$$

This algorithm is of square-root type since $\boldsymbol{U}\boldsymbol{D}^{1/2}$ is a covariance matrix square root. '1 'he numerical stability and accuracy of this algorithm--which relates the U-D observation update cycle to the numerically stable Givens orthogonal transformation method- has been established by Gentleman (1 973). Moreover, the U-D algorithm approaches the conventional KF in both timing and storage requirements if coded efficiently, e.g., exploiting the zero lower-triangular portion of the $\boldsymbol{U}$ factor in Eq. (2b) by storing and performing computat ions on nonzero matrix entries only.

U-D factorization is applied to both the covariance propagation in time (forecast) and data update (assimilation) in Kalman's sequential algorithm, so that all elements of the discrete KF equations- including the gain matrix--are expressed in terms of these factors. Wc outline here only the U-D form of the covariance matrix propagation step, since it is the dominant computational burden in the KF. The KF covariance propagation equation, in the notation of Ghil and Malanotte-Rizzoli (1 991), is

$$\boldsymbol{P}_k^f = \boldsymbol{\Psi}_{k-1}\boldsymbol{P}_{k-1}^a\boldsymbol{\Psi}_{k-1}^T + \boldsymbol{Q}_{k-1}, \tag{3}$$

where $\boldsymbol{P}_k^f$ and $\boldsymbol{P}_{k-1}^a$ arc the forecast covariance at time step k and the analysis covariance at step k-1, and Y' and $\boldsymbol{Q}$ are the state-transition and process-noise (model error) matrices. Using Eq. (1), Eq. (3) can be written as

$$\boldsymbol{U}_k^f\boldsymbol{D}_k^f\boldsymbol{U}_k^{fT} = \boldsymbol{\Psi}_{k-1}\boldsymbol{U}_{k-1}^a\boldsymbol{D}_{k-1}^a\boldsymbol{U}_{k-1}^{aT}\boldsymbol{\Psi}_{k-1}^T + \boldsymbol{\Lambda}_{k-1}\boldsymbol{Q}_{k-1}'\boldsymbol{\Lambda}_{k-1}^T, \tag{4}$$

with A defined so that $\boldsymbol{Q}'$ is diagonal. The right-hand side can be factored as $\boldsymbol{Y}\boldsymbol{Z}\boldsymbol{Y}^T$, where

$$\boldsymbol{Y} = \left[\boldsymbol{\Psi}_{k-1}\boldsymbol{U}_{k-1}^a \;\vdots\; \boldsymbol{\Lambda}_{k-1}\right], \qquad \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{D}_{k-1}^a & \vdots & \boldsymbol{0} \\ \cdots & \cdots & \cdots \\ \boldsymbol{0} & \vdots & \boldsymbol{Q}_{k-1}' \end{bmatrix}. \tag{5a,b}$$

This factorization has the same general form as Eq. (1), but not the same dimensions, and Y is no longer upper triangular. The cycle is completed by transforming Y and $\boldsymbol{Z}$ to the desired (n x n) unitary upper triangular and diagonal form, using a weighted Gram-Schmidt orthogonalization procedure (Golub and Van Loan 1989, Sec. 5.2.8) on $\boldsymbol{y}_i$, the rows of Y, that generates a set of n $\boldsymbol{Z}$-orthogonal vectors, $\boldsymbol{b}_i$. The elements of the desired $\boldsymbol{D}_k^f$ and $\boldsymbol{U}_k^f$ arc then given by

$$d_{jj} = \boldsymbol{b}_j^T\boldsymbol{Z}\boldsymbol{b}_j, \quad j = 1, \cdots, n; \qquad u_{ij} = \begin{cases} (\boldsymbol{y}_i^T\boldsymbol{Z}\boldsymbol{b}_j)/d_{jj}, & i = 1, 2, \cdots, j-1, \\ 0, & i > j, \end{cases} \tag{6a,b}$$

which completes the propagation from $\boldsymbol{D}_{k-1}^a$ and $\boldsymbol{U}_{k-1}^a$ to $\boldsymbol{D}_k^f$ and $\boldsymbol{U}_k^f$.

A measure of the potential for computational difficulty is the condition number, $\kappa(\boldsymbol{P})$, defined as $\lambda_{max}/\lambda_{min}$, where λ_{max} and λ_{min} arc the maximum and minimum eigenvalues of the symmetric, positive-definite matrix $\boldsymbol{P}$. If K(P) is large---on the order of 10^r, where r is the number of significant digits in a computer word—computational difficulties are certain to arise. Conditioning is improved by using estimation equations that usc the square roots of $\boldsymbol{P}$ rather than $\boldsymbol{P}$ itself, since for the square-root matrix $\kappa(\boldsymbol{U}\boldsymbol{D}^{1/2}) = 10^{r/2}$.

To introduce sparseness into the discrete estimation equations, wc approximate the full $\boldsymbol{P}$ by the symmetric banded approximation $\boldsymbol{P}_B$:

$$\boldsymbol{P}_B = \begin{bmatrix} \overbrace{\begin{matrix} x & x & \cdots & x \\ x & x & x & \cdots & x \\ \vdots & x & x & x & \cdots & x \\ x & x & x & x & x & \cdots & x \\ & x & x & x & x & x & \cdots & x \end{matrix}}^{\beta} & & \boldsymbol{0} \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \boldsymbol{0} & & & x & x & \cdots & x & x & x \\ & & & & x & x & \cdots & x & x \\ & & & & & x & x & \cdots & x \\ & & & & & & & x & x \; . \; . \end{bmatrix} \tag{7}$$

where the width of the central band in terms of nonzero diagonals is $2\beta+1$; $\beta = 3N_{lon}(2b+1)$, with N_{lon} the number of longitudinal grids points, and b the bandwidth parameter given by the number of grid points—away from a base point in each coordinate direction—over which nonzero covariances will be calculated. Only the nonzero diagonals of P_B arc stored. For P_B to be decomposable into U-D form, it cannot preserve any zero diagonals imbedded in the central band. Thus, our central diagonal band is *a priori* completely nonzero, while PC'S was interspersed with sub-bands of zero diagonals, due to block-bandedness. This kcy difference allows us to decompose. the sparse approximation P_B into U-D square-root factors: if P_B has the form (7), then $P_B = U_B D U_B^T$, where U_B is unit upper triangular and D is diagonal, if and only if U_B has the form

$$U_B = \begin{bmatrix} 1 & \overbrace{x \quad x \quad \cdots \quad x}^{\beta} & & & & 0 \\ & 1 & x \quad x \quad \cdots \quad x & & & \\ & & 1 \quad x \quad x \quad \cdots \quad x & & & \\ & & & \ddots & & \\ & & & 1 \quad x \quad x \quad \cdots & & x \\ & & & & 1 \quad x \quad \cdots & x \\ & 0 & & & \ddots & \vdots \\ & & & & 1 & x \\ & & & & & 1 \end{bmatrix} \tag{8}$$

The banded covariance approximation (7, 8) is more computationally expensive (i.e., with more nonzero elements needing to be stored and computed) than that of PC, but it allows us to retain the U-D factorized form, which, in turn, yields a near-optimal discrete filter that doesn't diverge. The new U-D square-root filter still takes full advantage of the induced sparseness by not storing— or doing computations on—the $n-\beta-1$ zero diagonals in the upper-right portion of the U_B factor (8).

3. The nonlinear shallow-water model

The model used in this study is a 2-D nonlinear shallow-water model, which describes divergent barotropic motion in a hydrostatic fluid with a free surface, confined to a beta-plane channel centered at 45"N (Todling and Ghil, 1994). The discretization is based on a second-order accurate, cluaclratic-energy -conserving finite-difference scheme (e.g., Grammeltvedt 1969), using a modified Euler-backward method in time. The prognostic variables are the zonal and meridional velocities, u and v, and geopotential height h at each gridpoint. The boundary conditions are very simple: v is set to zero at the meridional boundaries of the computational domain, and u and h are independent of longitude along the same boundaries. All fields arc periodic in the zonal direction.

For estimation purposes, the dependent variables arc organized into an ordered state vector of length n, where n is three times the number of grid points. The nonzero diagonals of the sparse $n \times n$ state transition matrix Y' arc computed analytically with a tangent-linear approximation, This matrix is used in advancing P_k^f as in Eq. (3), while the state itself is advanced by the full nonlinear equations of motion.

4. Numerical results

4.1 *U-D filter test*

We have run several assimilation experiments on a 17 x 24 grid with boundaries at 25°N and 65"N (5000 km meridionally), so that Δx = *1178* km and Δy = 312.5 km, with an integration step size of Δt = 15 *rein*, and with n = 3 x 17 x 24 = 1224 state variables. This model size is comparable to advanced linear KF implementations (Jiang and Ghil 1993, Miller *et al.* 1994, Todling and Ghil 1994), but the present method should permit, once tested, much larger applications. The height h of the model's free surface is given initially, and the initial data for the velocity components u and v are calculated from the geostrophic relation. The initial state is a westerly jet with north-south perturbations of different wavelengths and amplitudes along its zonal axis. Synthetic observations are taken every 3 hours, and arc processed using the U-D square-root EKF of Sec. 2 with the bandwidth parameter b ranging from 2 to 4 and full (i.e., no approximation, $\beta = n$ = 1224). A diagonal model-error (process-noise) matrix Q with nonzero variances for all variables was used in all the experiments.

For the relatively small model size used in these experiments, the banded approximation yields modest computational savings. For example, for b = 3, the number of nonzero diagonals of the covariance matrix that must be stored and calculated is 504 out of a total of n = 1224. This yields a savings of about 1/2 in the floating-point operation count, since the centrally located diagonals retained are longer than the ones omitted. However, for a more realistically sized atmospheric or oceanic model, the savings can be more dramatic: for a global model with a 2° x 2° grid using a bandwidth parameter b = 4, only 10% of the diagonals of the full covariance matrix arc retained and calculated with the approximation. As an accuracy check of the banded U-D filter, we note that the resulting state estimates and their associated error-covariance estimates- obtained using b = 2, 3, and 4— did not differ in any significant way from those obtained using the full U-D filter.

On the other hand, a check of the U-D filter's overall performance relative to that of a conventional EKF (CKF hereafter) yielded dramatic results. Every U-D filter experiment was successful in that the filter performed in a numerically stable, ncm-divergent manner; the same cases run with the CKF- on the same CRAY YMP using the same precision—invariably failed duc to filter divergence after a few forecast-update steps. The failures were all similar: each would terminate due to a sudden rapid increase in the norm of the computed P_k^f matrix, culminating in computer overflow within one day after initial time. It is interesting that compari son CKF runs replicated almost exactly the full U-D filter results ($b = n$) during the first few assimilation steps; soon, however, results would begin to drift apart duc to the accrual of numerical errors in the unstable CKF computations, followed rapidly by the CKF's blow-up.

4.2 *State and covariance estimates*

Wc show the results of an "identical-twin" experiment where, starting with an initial "guess" $w^a(t_o)$ of the "true" initial state $w^t(t_o)$, wc use noisy observations of the "true" evolution $w^t(t), t > t_0$, to obtain a U-D filter estimate $w^a(t)$ that converges to the true field $w^t(t)$ as assimilation proceeds. The initial error $w^a(t_o) - w^t(t_o)$ was generated with a geostrophically balanced perturbation that has approximate root-mean-square (RMS) values of $\Delta u = 3.8$ m/s and $\Delta v = 1.29$ m/s for the velocities, and $\Delta h = 370$ m for the geopotential height (Fig. 1), with the RMS taken over all grid points. The observing pattern used was highly idealized and the observations rather plentiful in this preliminary experiment: observations of u, v, and h were made c-very 12 At = 3 hours at all grid points.

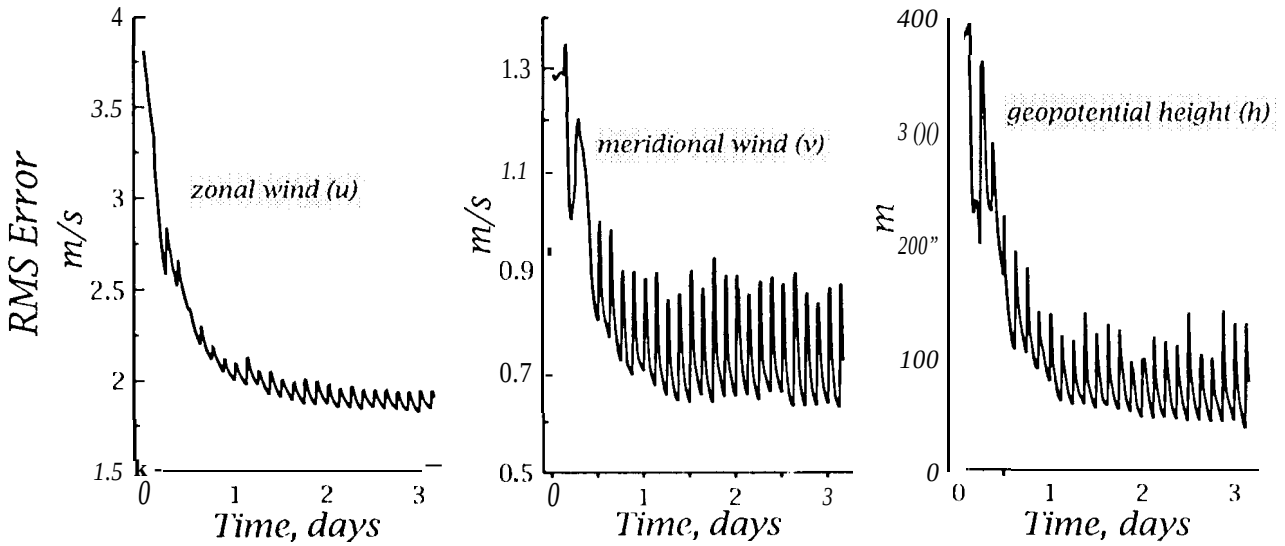

Figure 1. The 3-day evolution of the RMS assimilation error taken over all the grid points of the beta-plane channel. The estimate error for the wind components u and v and the geopotential height h are shown individually at each forecast step. There arc 96 forecast steps pcr day, and the fields are updated every 12 At (3 hours).

Observations were generated by adding random, uncorrelated Gaussian noise to the true trajectory, with standard deviations $\sigma_u = 2$ m/s, $\sigma_v = 0.75$ m/s, and $\sigma_h = 75$ m. The observations were processed sequentially at each assimilation time (Ghil and Malanotte-Rizzoli 1991, p. 185, and references there), with each row of the observation matrix H containing a single entry equal to one. and the rest zeros. Model e. I rors were also assumed to be uncorrelated, with standard deviations of 10 m/s, 5 m/s, and 200 m pcr half-day for u, v, and h, respectively. The initial analysis-error covariance matrix, P_o^a, was taken to be diagonal, with standard deviat ions given by 2S rids, 8 m/s, and 1000 m for u, v, and h. Figure 1 shows the. RMS error results of a 3-day assimilation run at each forecast and analysis step: the RMS errors in the u, v, and h fields are computed from the vector $w^f(t) - w^t(t)$ by summing over all grid points. The estimated state is seen to converge to the true state in all components.

Figure 2 depicts the evolution of the estimated error in u. Each panel in the figure shows the state-estimate error at a specific forecast time over the rectangular beta-channel region with a gray-scale range depicting errors in the wind component ranging from -4 m/s (black) to +4 m/s (white). The images progress in time from 6 **hours to 48** hours past initial time, and show the estimated state converging over the entire spatial domain to the true trajectory.

A closer inspection of Fig. 1 reveals a manifestation of a problem that will be. at once: (1) new to many KF practitioners; (2) explainable only with advanced numerical analysis concepts; (3) difficult, but not impossible, to solve; :irld----asidc from the computational burden- (4) probably at the heart of what will be the major difficulty in applying the EKF to data assimilation for highly resolved geophysical flow models. In Fig.], at the update time itself, the state estimation error *increases*, rather than decreases. This problem arises from the fact that errors in the observational residual-the difference between an observed value and the. model's prediction of that value- arc transformed by the Kalman inversion process, even in the square-root form, into large errors in the. state estimate along those directions in the

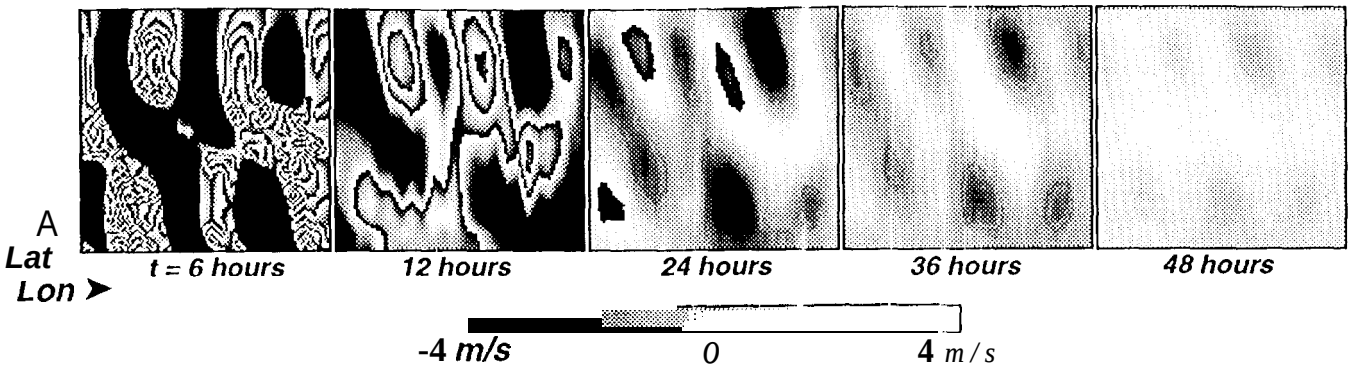

Figure 2. Time evolution of the estimated state error in u; gray scale. 1 Each image depicts the error over the rectangular beta-channel region which constitutes the domain of the cm-layer shallow-water model used. The 17 parallels run 25° to 65° N, and the 24 meridians run from 0° to 360° E.

r?-dimensional phase space that correspond to eigenvectors of P_k^f associated with large eigenvalues. A heuristic explanation of this fact in the context of large-scale geophysical, mid-latitude flows, is that the divergence of the flow field is much smaller that the curl. Hence, errors in observed divergence will affect the assimilated result much more seriously than errors of similar size in the curl. This problem is ubiquitous in EKF estimates of large nonlinear systems that (1) require linearized approximations, and (2) possess covariance matrices- and hence their inverses, the system information; matrices- that are ill-conditioned.

Figure 3 shows the evolution in time of the spectrum: each curve is the]]lagilitllclc-ordered set of n=1224 eigenvalues computed from the forecast-error covariance matrix that results at each of 6 forecast steps, namely steps 1, 2, 3, 4, 13, and 37 following initial lime t_0, with the updates occurring after steps 13, 25, 37, etc. A significant increase, of almost 10^6, in the condition number is evident in the early filtering stages; this implies a serious numerical stability problem in the making. The system that must be (at least implicitly) inverted to proceed with the KF process is seen to be rapidly reaching the stage that standard inversion methods will not work.

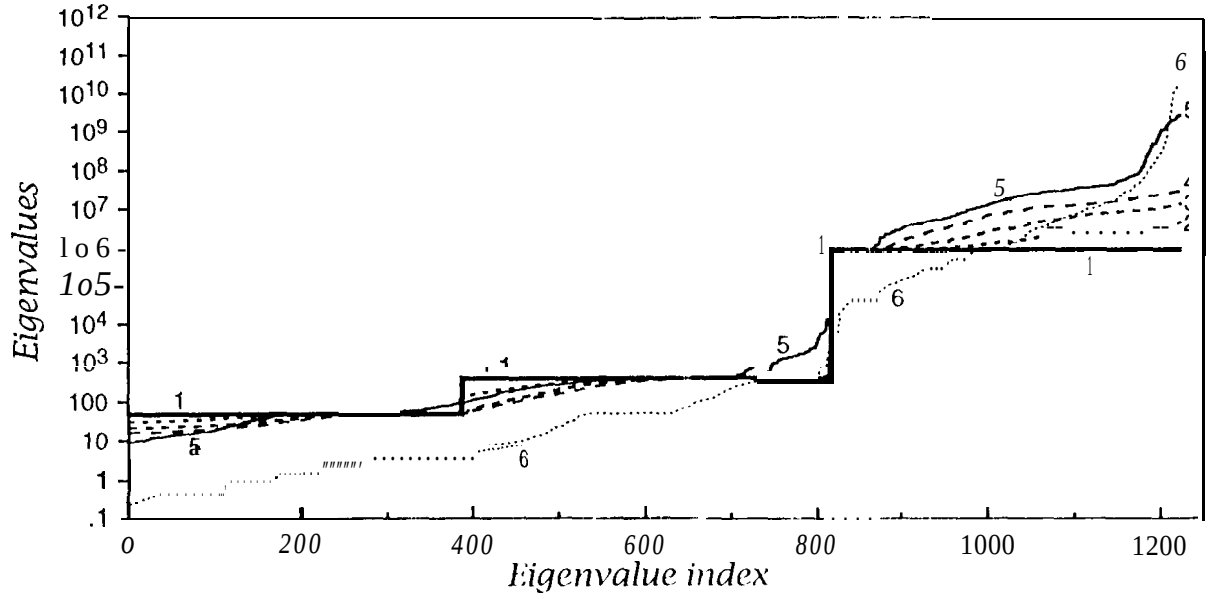

Figure 3. Magnitudc-ordered spectrum of the 12.24 x 1224 forecast-error covariancc matrix at 6 different time steps of the assimilation. Curve 1 (bold) shows the. eigenvalues for the 3-tiered initial *a priori* uncertainty for each of the 3 variables u, v, and h. Curves 2, 3, and 4 arc the eigenvalues for forecast steps 2, 3, 4, immediately following the initial time. Curve 5 results at the last forecast step immediately preceding the first analysis time. Curve 6 shows the eigenvalues preceding the third analysis time (step 37).

5. Summary

We have shown that a square-mot implementation of the EKF for a fairly realistic shallow-water model is stable. This can overcome the stability problems of the CKF encountered in space-navigation problems, many other engineering areas (Kerr 1990, and references there), by Evensen (1992) in an ocean model, and in the present atmospheric model.

Using U-D factorization, a modified version of Parrish and Cohn's **(1985)** banded approximation is stable. Until the CKF begins to diverge, it agrees quite well with the banded U-D square-root filter.

A square-root EKF has the best chance of dealing successfully with the problems posed by ill-conditioning of large error-covariance matrices. The difficulties still encountered with the present U-D implementation will be surmounted in the future by changing to a square-root information filter (SRIF: Bierman 1977).

6. References

Balgovind, R., A. Dalcher, M. Ghil and E. Kalnay, "A stochastic dynamic model for the spatial structure of forecast error statistics," *Mon. Wea. Rev., 11S,* 273-296, 1983.

Bellantoni, J. F. and K. W. Dodge, "Square-root formulation of the Kalman-Schmidt filter," *AIAA J., 5, 1309-1314,* 1967.

Bier man, G. J., "*Factorization Methods for Discrete Sequential Estimation,*" Academic Press, New York, N. Y., 1977.

Dyer, P. and S. McReynolds, "Extension of square-root filtering to include process noise," *J. Opt. Theory Applic.*, **3**, *444-458, 1969.*

Evensen, G., "Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model," *J. Geophys. Res.,* 97, 17,905 --J 7,924, 1992.

Gentleman, W. M., "Least squares computations by Givens transformations without square roots," *J. Inst. Math. Appl.*, 12,329-336, 1973.

Ghil, M. and P. Malanotte-Rizzoli, "Data assimilation in meteorology and oceanography," *Advances in Geophysics,* 33, Academic Press, 141-266, 1991.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube anti E. Isaacson, "Applications of estimation theory to numerical weather prediction," In: *Dynamic Meteorology: Data Assimilation Methods, L.* Bengtsson, M. Ghil, and E. Kallen, eds., 139-224, 1981.

Golub, G. and C. Van Loan, *"Matrix Computations,"* 2nd cd. John Hopkins University Press, Baltimore, Md., 1989,

Grammeltvedt, A., "A survey of finite-diffcrmcc schemes for the primitive equations for a barotropic fluid," *Mon. Wea. Rev.,* 97,384--404, 1969

Jazwinski, A. H., "*Stochastic Processes and Filtering Theory,*" Academic Press, Ncw York, N. Y., 1970.

Jiang, S. and M. Ghil, "Dynamical properties of error statistics in a shallow-water model," *J. Phys. Oceanogr.,* 23, 2541--2566, 1993.

Kalman, R. E., "A new approach to linear filtering and prediction problems," *Trans. ASME. (J. Basic Eng.),* Vol. **82D**, 34-45, 1960.

Kerr, T. H., "Fallacies in computational testing of matrix positive definiteness/semidefiniteness," *IEEE Trans. Aero. Elect. Systems,* 26,415-421, 1990.

Lawson, C. and R. Hanson, "*Solving Least Squares Problems,*" Prentice-Hall, Englewood Cliffs, N. J., 1974.

Leondes, C. T. (Ed.), "Theory and applications of Kalman filtering," NATO Advisory Group for Aerospace Res. Develop., *AGARDograph* 139, *Rep.* No. *AD* 704.? 06, 1970.

Miller, R. N., A. J. Busalacchi and E. C. Hackert, "Sea surface topography fields of the tropical Pacific from data assimilation," *J. Geophys. Res.,* in press, 1994.

Parrish, D. F. and S. E. Cohn, "A Kalman filter for a twc)-dimensional shallow water model: formulation and preliminary experiments," *Office Note 304,* NOAA-National Meteorological Center, Washington, DC. 20233, 1985.

Schlee, F. H., C. J. Standish and N. F. Toda, "Divergence in the Kalman filter," *AIAA J., 5, 1114-1120,* 1967.

Schmidt, S. F., J. D. Weinberg and J. S. Lukesh, "Case study of Kalman Filtering in the C-5 aircraft navigation system," 1968 *Joint Automatic Control Conf., Case Studies in System Control.*, 1968.

Thornton, C. L. and G. J. Bierman, "A numerical comparison of discrete Kalman filtering algorithms: an orbit determination case study," *Jet Propulsion Laboratory Technical Memorandum TM* .7.7 -771, 1976.

Todling, R. and M. Ghil, "Kalman filtering for a two-layer, two-dimensional shallow-water model," in *Proc. Intl. Symp. Assimilation Obsns. Met. Oceanogr.*, Clermont-Ferrand, 454-459, 1990.

Todling, R. and M. Ghil, "Tracking atmospheric instabilities with the Kalman. Part I: Methodology and one-layer results," *Mon. Wea. Rev., 122,* 183- *204, 1994.*